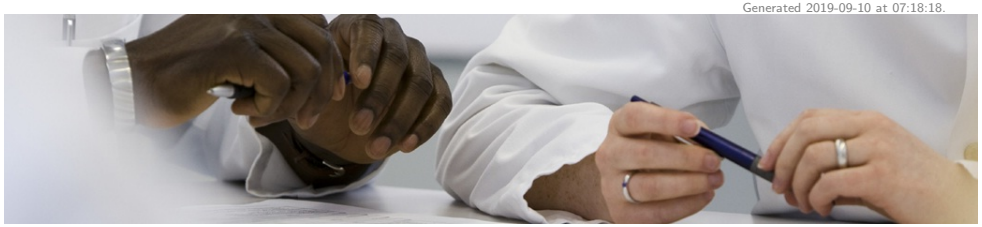
Before going BIG - seize opportunities for small data: multistate models to improve decision-making and effect quantification in clinical trials

Kaspar Rufibach
Methods, Collaboration & Outreach Group, Department of Biostatistics, Roche Basel
XXXIst Conference of the Austro-Swiss Region of the International Biometric Society
10h September 2019, Lausanne, Switzerland

Generated 2019-09-10 at 07:18:18.



Where

RESEARCH PAPER | Full Access |

A multistate model for early decision-making in oncology

Ulrich Beyer, David Dejardin, Matthias Meller, Kaspar Rufibach, Hans Ulrich Burger

First published: 16 July 2019 | https://doi.org/10.1002/bimj.201800250

Simulated dataset + comprehensive R code.

Who




How do we typically decide whether to move an oncology molecule into Phase 3?

Decision-making in early oncology development

1. Small single-arm trial for **experimental** drug (e.g. $n = 40$).
2. Response proportion, duration of response.
3. Compare to "corresponding" quantities from literature for **control** treatment.

But:

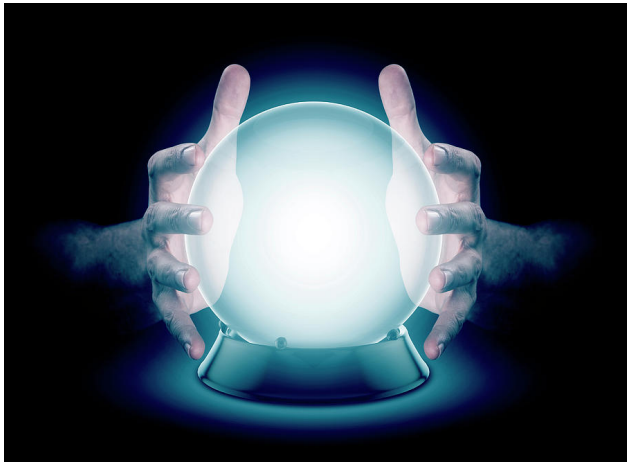
- **P(wrong decision)** may be high.
- Primary endpoint in Phase 3: **Overall survival**.

Proposal:

Decide in early phase based on OS prediction.

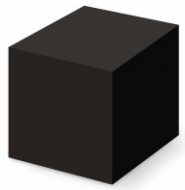
Decrease P(wrong decision).

Prediction?



Prediction?

Blackbox ML algorithm using big data?



Transparent multistate model with historical borrowing.

Challenges and proposal

Challenges:

1. Response-type endpoint?
2. Surrogacy? **Poor** in many indications.
3. Immunotherapy (CIT): no effect on response, relevant OS effect.
4. **Non-randomized** comparison $\Rightarrow$ confounding.

Proposal: Base decision-making on **OS prediction from multistate model**.

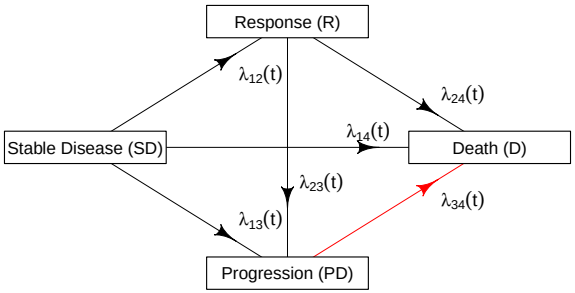
1. **Predicted survival function for experimental arm**.
2. Combine S_{exp} with S_{control} to get **predicted OS HR**.
3. Experimental drug might act on certain transitions only $\Rightarrow$ not captured through simple modelling of OS. Potential **efficiency gain**!
4. **Propensity scoring**.

Idealized scenario: Retrospective data from Phase 3 RCTs.

Long-term follow-up in both arms.

Randomization $\Rightarrow$ no confounding.

Multistate model for early decision-making



- Follow-up of patient until **PD or death without PD**.
- Post-progression hazard λ_{34}: **borrowing** from historical data.
- Transitions SD $\rightarrow$ D, R $\rightarrow$ D rare, hazards $\approx$ same in both arms.

Predicted survival function in experimental arm, S_{exp}

Compute transition probabilities for each transition.

$$S_{\text{exp}}(t) = 1 - \Big(P_{SD\rightarrow D}(0,t) + P_{SD\rightarrow \text{PD}\rightarrow \text{D}}(0,t) + P_{SD\rightarrow R\rightarrow D}(0,t) + P_{SD\rightarrow R\rightarrow \text{PD}\rightarrow \text{D}}(0,t)\Big).$$

λ_{34} corresponding to **PD $\rightarrow$ D** transition borrowed from historical data.

Historical borrowing for λ_{34}

Experimental treatment expected to provide benefit **beyond PD**?

No:

- E.g. chemotherapy or antibody-dependent cellular cytotoxicity.
- **Plug-in** hazard function estimate from historical control.
- No post-PD information required for experimental arm.

Yes:

- E.g. chemoimmunotherapy.
- Estimate post-PD hazard ratio assuming **proportionality**.
- How much post-PD deaths needed in experimental arm to reliably **estimate post-PD HR**?

Benefit beyond PD: Oak

Oak

Previously treated non-small-cell lung cancer.

Rittmeyer et al. (2017).

	Atezolizumab	Chemotherapy	Hazard ratio
Effect post-PD	expected	not expected	
Objective Response	58 (13.6%)	57 (13.4%)	
Duration of Response	26.3 (10 - ∞)	6.2 (4.9 - 7.6)	
Overall Survival			0.73 (0.62, 0.87)

If this were early phase data - would you initiate Phase 3?

Competitors used this mechanism of action.

OS prediction when post-PD hazards assumed proportional

Random variable:

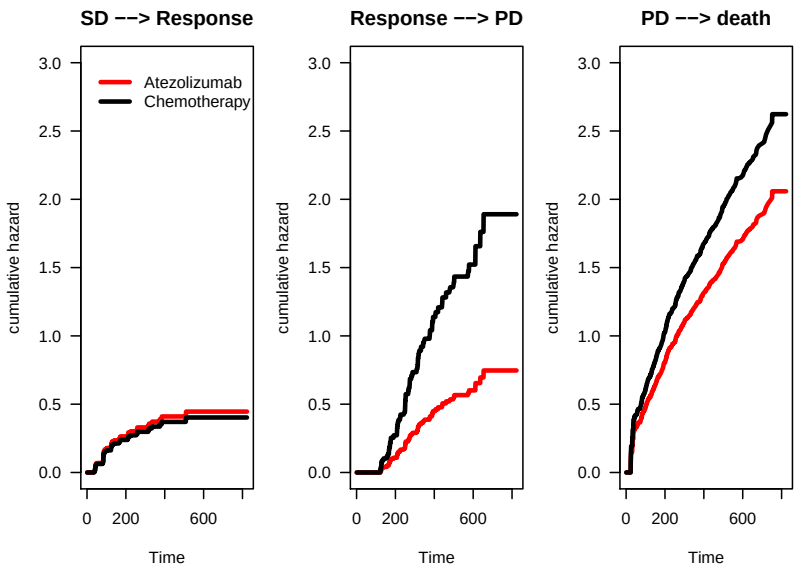
$$Z = \begin{cases} 0 & \text{if patient in control,} \\ 1 & \text{if in experimental group.} \end{cases}$$

$$\lambda_{34}(t \mid Z) = \lambda_{34,0}(t) \exp(\beta_{34} Z)$$

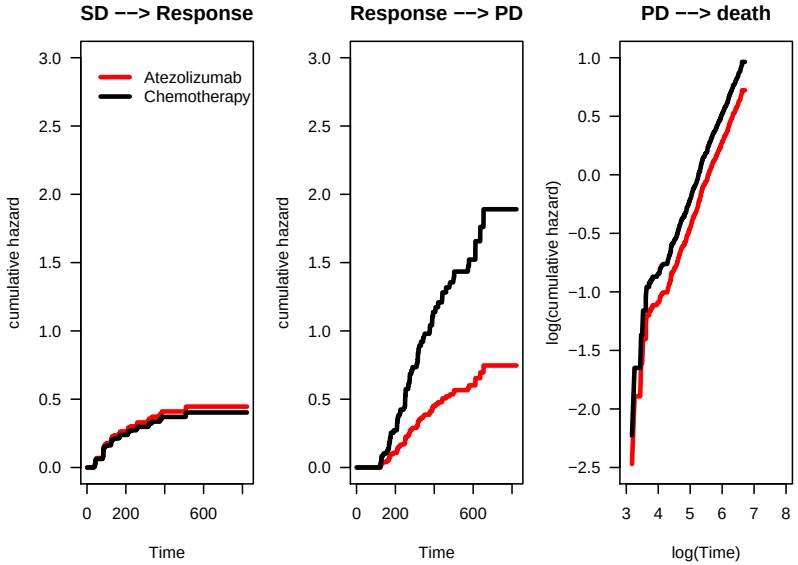
Baseline hazard $\lambda_{34,0}$ **estimated from both arms combined**.

Post-progression hazard ratio β_{34}?

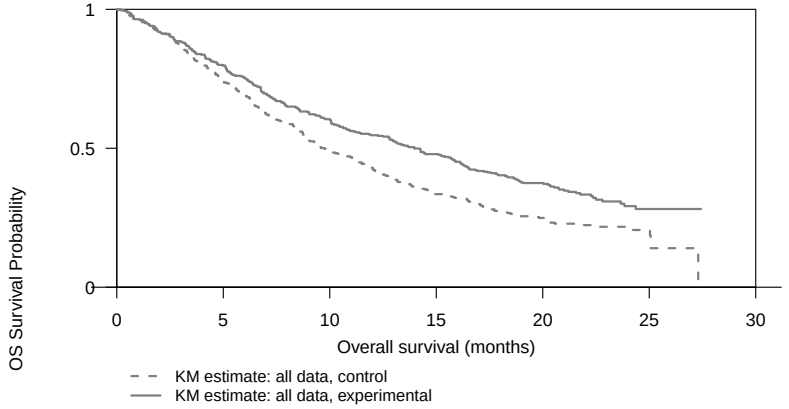
Oak: raw cumulative hazard estimates (of interest)



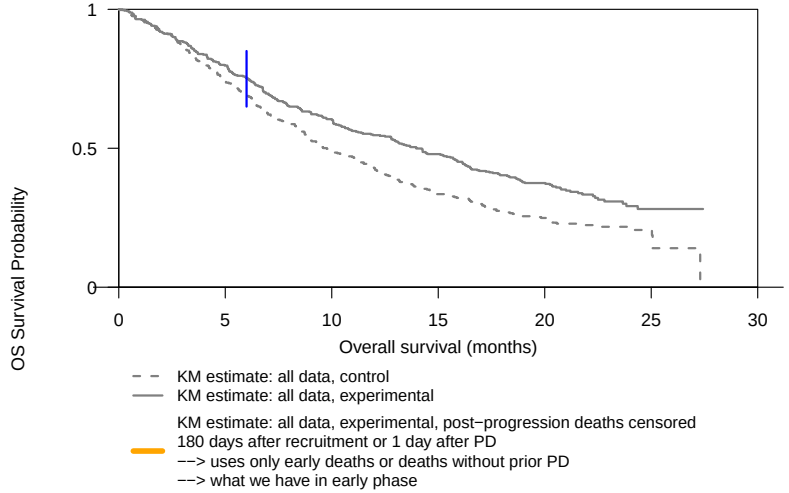
Oak: raw cumulative hazard estimates (of interest)



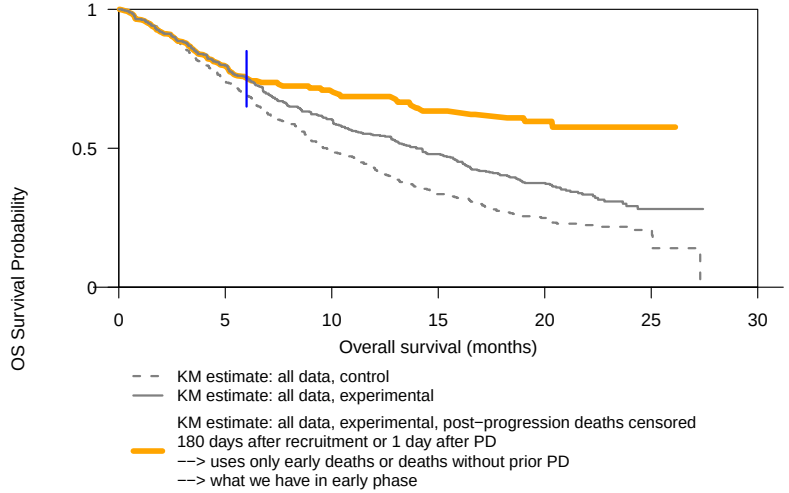
Oak: estimates / predictions of S_{exp}



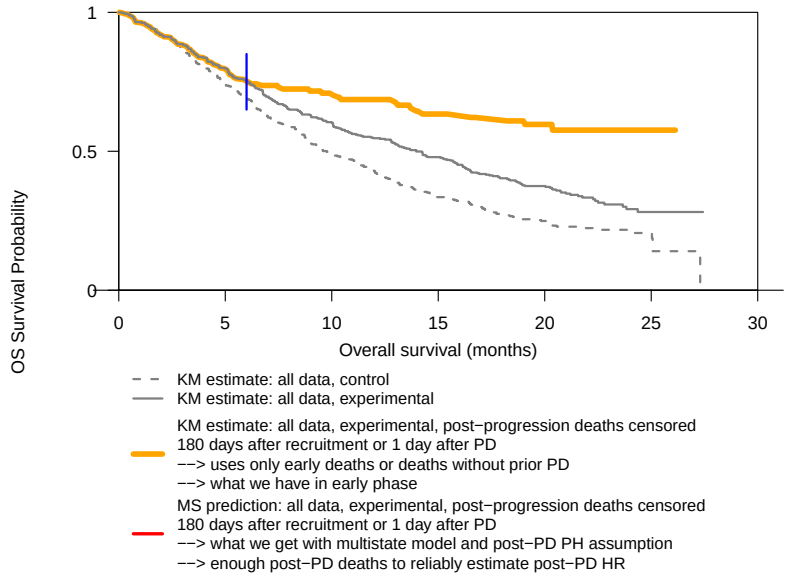
Oak: estimates / predictions of S_{exp}



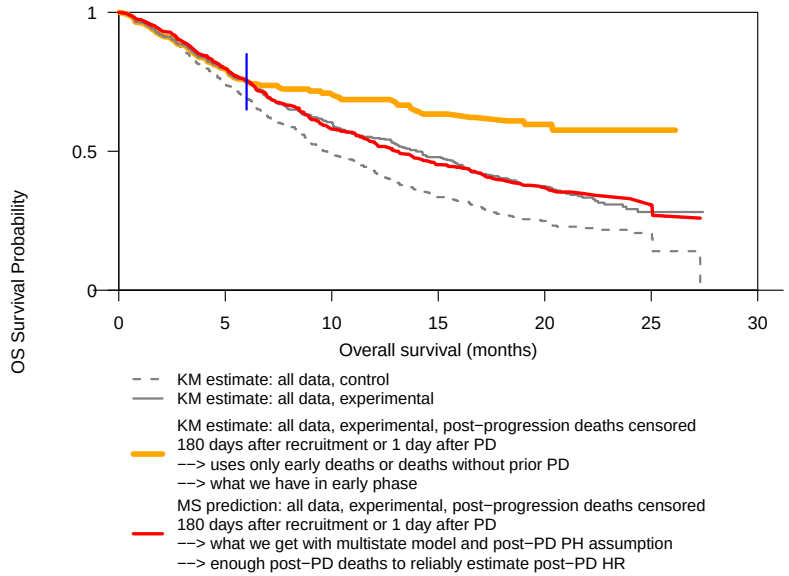
Oak: estimates / predictions of S_{exp}



Oak: estimates / predictions of S_{exp}



Oak: estimates / predictions of S_{exp}



Early phase decision based on multistate prediction:

P(wrong decision)?

OS prediction from mimicked early phase data

Historical control: Oak control arm data.

False-positive decision: Sample early phase trial from Oak control arm.
False-negative decision: Sample early phase trial from Oak experimental arm.

Sample early phase trial:

- 40 patients,
- 6 months uniform recruitment,
- analysis 15 months after first patient entered,
- censor post-PD follow-up **one day after PD**,
- estimate $\lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{23}, \lambda_{24}$ from this data.

Cox regression for post-PD transition $\Rightarrow \widehat{\lambda}_{34}(t|Z)$.

Compute prediction of S_{exp}.

OS HR prediction based on early phase trial

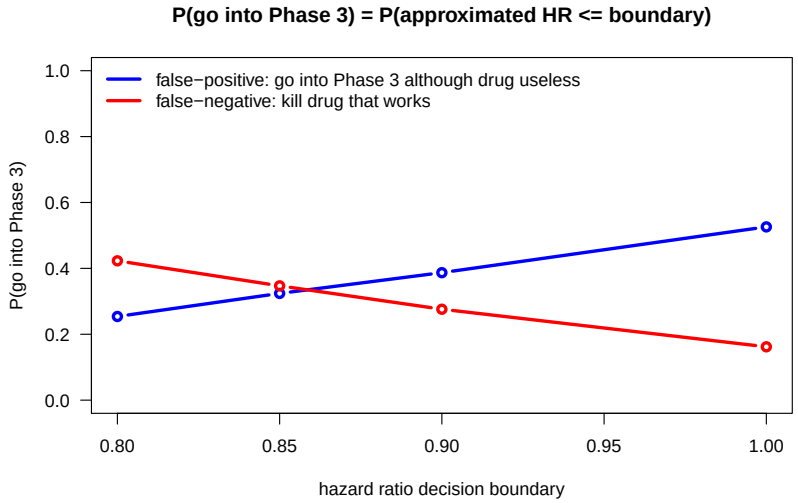
Approximate HR by fitting exponential distribution to both arms $\Rightarrow \widehat{HR}$.

Decision to move to Phase 3: $\widehat{HR} \leq$ boundary $\in \{0.80, 0.85, 0.90, 1.00\}$.

Repeat 1000 times.

Resampling $\Rightarrow$ **quantification of uncertainty**.

Oak: P(wrong decision)



How many post-PD deaths to estimate HR of PD $\rightarrow$ death transition?

How many post-PD deaths to estimate HR of PD $\rightarrow$ death transition?

Ask during Q&A.

Conclusions for early-decision making proposal

Conclusions

Early phase decision-making based on **multistate OS prediction**:

- Assumption on $\lambda_{34} \Rightarrow$ need to understand **disease and treatment**.
- **Avoids difficulty in interpretation of response-type endpoints**.
- Feasibility assessed in **idealized scenario**.
- Recommendation **how much post-PD follow-up** needed to estimate β_{34}.
- Needs **long-term individual-patient** data in control arm!

What about confounding?

Real-world data as historical control.

Combine proposal with propensity scoring.

Immunotherapy:
1) no difference in PFS,
2) non-proportional hazards for OS.

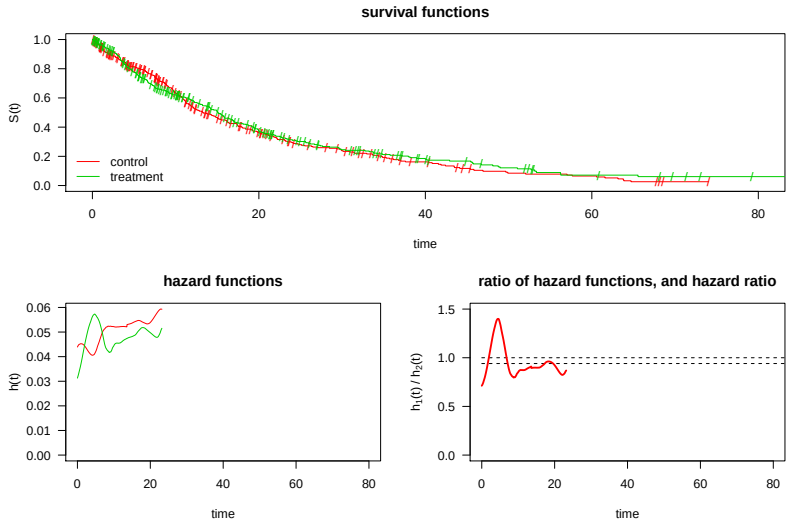
How to quantify effect?

A fictional clinical trial

Simulated clinical trial:

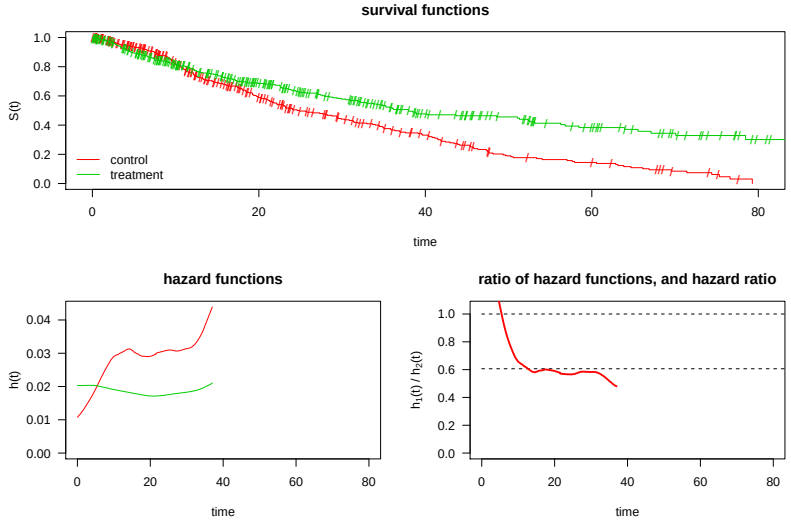
- 1:1 randomized, 400 and 400 patients per arm.
- No administrative censoring, but drop-out.

PFS for simulated clinical trial



- Estimated hazard ratio: 0.94, 95% confidence interval [0.80, 1.11].
- Hypothesis test for PH: $p = 0.24$.

OS for simulated clinical trial



- Estimated hazard ratio: 0.61, 95% confidence interval [0.50, 0.74].
- Hypothesis test for PH: $p < 0.0001$.

Summarize treatment effect

Non-proportional hazards for OS. How to summarize effect of treatment?

Data was generated according to:

Transition	Control arm	Treatment arm
$0 \to 1$	$\lambda_{01}^c = \log(2)/25$	$\lambda_{01}^t = \lambda_{01}^c \cdot \mathbf{1}$
$0 \to 2$	$\lambda_{02}^c = \log(2)/30$	$\lambda_{02}^t = \lambda_{02}^c \cdot \mathbf{0.8}$
$1 \to 2$	$\lambda_{12}^c = \log(2)/15$	$\lambda_{12}^t = \lambda_{12}^c \cdot \mathbf{0.4}$

	coef	HR = exp(coef)	95% CI	p-value
transition event-free –> PD	-0.04	0.96	[0.77, 1.19]	0.72
transition event-free –> death	-0.09	0.91	[0.70, 1.18]	0.49
transition PD –> death	-1.09	0.34	[0.24, 0.46]	< 0.0001

Gaschler-Markefski et al. (2014).

Conclusions

Multistate models

Multistate models useful:

- Canonical **extension of survival analysis**.
- Get more **insight** in how disease and drug work.
- **Prediction** in well-specified, as opposed to black-box, model.
- **Jointly** model three key oncology endpoints: response, PFS, OS.
- Applications by no means restricted to oncology!

Many potential applications:

- Improved **early stage decision-making** ⇒ Beyer et al. (2019).
- Improved **communication** of effect and optimized **sample size** computation.
- Bivariate modelling of PFS and OS to help inform **surrogacy** questions ⇒ Meller et al. (2019).

Big vs. small data

Often, information removed/altered in **small data**:

- (Artificial) response cateogries instead of actual measurements: **dichotomization**,
- response proportions only: **ignoring the dynamics between states**,
- complicated subsets, e.g. those that respond only: **selection bias**,
- effect quantification in **one** number where biological process might suggest few numbers,
- ...

Maximize information from **small** data. AND look at **BIG** data.

Biostatisticians ideally placed to contribute to this!

Thank you for your attention

References I

- Andersen, P. K., Borgan, r., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer.
- Baselga, J. and Cortes, J. e. a. (2012). Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *N. Engl. J. Med.* **366** 109–119.
- Beyer, U., Dejardin, D., Meller, M., Rufibach, K. and Burger, H. U. (2019). A multistate model for early decision making in oncology. *Biom J, to appear* .
- Gaschler-Markefski, B., Schiefele, K., Hocke, J. and Fleischer, F. (2014). *Multi-state Models Used in Oncology Trials*. Springer Berlin Heidelberg, Berlin, Heidelberg, 283–304. `https://doi.org/10.1007/978-3-642-55345-5_16`
- Meller, M., Beyersmann, J. and Rufibach, K. (2019). Joint modelling of progression-free and overall survival and computation of correlation measures. *Stat. Med.* **38** 4270–4289.
- Rittmeyer, A., Barlesi, F., Waterkamp, D., Park, K., Ciardiello, F., Von Pawel, J., Gadgeel, S. M., Hida, T., Kowalski, D. M., Dols, M. C. et al. (2017). Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (oak): a phase 3, open-label, multicentre randomised controlled trial. *The Lancet* **389** 255–265.
- Swain, S. M. and Baselga, J. (2015). Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer. *N. Engl. J. Med.* **372** 724–734.

Doing now what patients need next

R version and packages used to generate these slides:
R version: R version 3.6.0 (2019-04-26)
Base packages: stats / graphics / grDevices / utils / datasets / methods / base
Other packages: nls2 / proto / diagram / shape / ggplot2 / rocheBCE / muhaz / flexsurv / reporttools / xtable / mstate / etm / dplyr / mvna / prodlim / biostatKR / survival

This document was generated on 2019-09-10 at 07:18:18.